Software Requirements Specification

for

FRAUD DETECTION SYSTEM.

Prepared by Tumusiime Arnold 17/U/10549/EVE, Muyanja Rodney 17/U/6819/EVE, Tusiime Allan 17/U/10630/EVE, Shaka Stephen 16/U/20180/EVE

2019-07-18

Table of Contents

Table of Contents ii

1. Introduction 1

1.1 Purpose 1

1.2 Document Conventions 1

1.3 Intended Audience and Reading Suggestions 1

1.4 Product Scope 1

2. Overall Description 1

2.1 Product Perspective 1

2.2 Product Functions 1

2.3 Operating Environment 2

2.4 Design and Implementation Constraints 2

3. NonFunctional Requirements 2

3.1 Performance Requirements 2

4. Use Cases For The Data Pipeline 2

4.1 Use Case Diagram 2

4.1.1 Description Of key Use Cases 2

4.1.2 The Use Case Diagram…………………………………………………………………...3

# Introduction

## Purpose

To implement a fraud detection system. This srs document contains all information about the implementation of the system.

## Document Conventions

This document uses the following conventions, Fontsize-14, Font-TimesNewRoman. It also uses bolded letters for headings of topics and subtopics.

## Intended Audience and Reading Suggestions

This document is intended for, data scientists, panelists, developers, project managers, account managers, finance officials, users, testers, and documentation writers. The rest of this SRS contains a description of the system, requirements, features, functionalities and references organized in the respective order. We Suggest the following sequence for reading the document, overview sections, requirements, features, functionalities and references.

## Product Scope

The project involves the implementation of the fraud detection system that will detect fraud within the loan accounts. The main goals of the system are to load the data about the loan accounts basing on their customer ids and base on that various data to detect the fraud in an account. The data will be used to perform analysis and create models from which one will be selected to detect fraud basing on the good-bad-flag.

# Overall Description

## Product Perspective

This product is a project assignment for our recess group and it has no other follow up state or earlier version

## Product Functions

* To load data from the csv file.
* To use the data to perform analysis by creating visuals.
* To use the data to create models for detecting the fraud.

## Operating Environment

The product will operate on a desktop computer or laptop computer, python anaconda IDE. The product will also operate on windows 10 operating system.

## Design and Implementation Constraints

Developers must be able to know how to use python 2.7 and python 3.0.

# Nonfunctional Requirements

## Performance Requirements

This product requires at least python 2.7 to be functional. It also requires windows 10 Operating System.

# Use Cases for the Data Pipeline.

## Use Case Diagram.

This is a visual summary of several related use cases within the data pipeline.

The use cases here include,

* Data Loading.
* Data wrangling.
* Data visualization.
* Data Modelling.
* Deployment.

The actors include,

* Data Scientist.

This actor uses the pipeline for loading data, cleaning the data, creating visuals and models using the loaded data. All the mentioned procedures have to be carried out so as to draw a meaningful conclusion of fraud from the data.

* Accounts Manager.

This actor uses the pipeline to view the analysis created from the data in visual form and also uses the results from the model so as to identify which accounts are being used for fraud.

### Description of key use cases.

* Data Loading. This simply involves loading data from a dataset into working environment.
* Data Wrangling. This involves, cleaning of data, label encoding, feature scaling, and removal of outliers.
* Data Visualization. This involves presentation of data analysis such as bar graphs, histograms, and so on, determining feature correlation and skewedness of data.
* Data Modelling. This involves creating models and training them to correlate the data with the outcomes.
* Deployment. This involves deployment of the selected model to be used in the fraud detection.

### The Use Case Diagram

Data Pipeline

Loads data

Cleans the data

Does data visualization

Data Scientist

Views the data visuals

Creates Model

Notifies Manager of the accounts with fraud

Deploys the model

Accounts Manager